

Spectroscopic and multivariate data-based method to assess the metabolomic fingerprint of Mediterranean plants

Laura Grauso

Maurizio Zotti

Wen Sun

Bruna de Falco

Virginia Lanzotti

Giuliano Bonanomi

This is the peer reviewed version of the following article:
Grauso, L. et al. (2019). Spectroscopic and multivariate data-based method to assess the metabolomic fingerprint of Mediterranean plants which has been published in final form at <https://doi.org/10.1002/pca.2862>

This article may be used for non-commercial purposes in accordance with [Wiley Terms and Conditions for Self-Archiving](#).

1 **Spectroscopic and multivariate data based method to assess the metabolomic**
2 **fingerprint of Mediterranean plants**

3

4 Laura Grauso^a, Maurizio Zotti^a, Wen Sun^a, Bruna de Falco^b, Virginia Lanzotti^a and Giuliano
5 Bonanomi^{a*}

6

7 ^a: Dipartimento di Agraria, Università di Napoli Federico II, Via Università 100, I-80055 Portici,
8 Naples, Italy

9 ^b: School of Science, Engineering & Technology, Division of Food & Drink, University of
10 Abertay, Bell Street, DD1 1HG Dundee, Scotland, UK

11

12 *To whom correspondence should be addressed. Tel.: +39 0812539015; fax: +39 081

13 E-mail address: giuliano.bonanomi@unina.it (G. Bonanomi).

14

15 Running Title: Metabolomic Fingerprint of Mediterranean Plants

16

17 **Keywords:** Metabolite profile, NMR, GC-MS, PCA, Metabolites, *Dittrichia viscosa*, *Hedera helix*.

18

19 **Short Abstract for Table of Content**

20 This study was aimed to a rapid and versatile assessment of the metabolomic fingerprint of
21 Mediterranean plants through an integrated approach of spectroscopic techniques, NMR and GC-
22 MS, and Multivariate Data Analysis. The chemical profile of eight Mediterranean species, largely
23 used as folk remedy, was analyzed. In addition, the metabolic compositions of leaves and roots
24 were compared for each species.

25

26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

ABSTRACT

Introduction – Most secondary metabolites from plants have a prominent defensive role and repellency against predators and microbial pathogens. These properties largely varies among plant species and found potential applications as biologically active compounds in medicine as well in agriculture.

Objectives – We propose a new procedure that combine different spectroscopic techniques and multivariate data analysis to determine the chemical composition and the relative amounts of each metabolites and/or each class of organic compounds. The approach was used for a rapid identification of secondary metabolites from leaf and root of eight Mediterranean plants species.

Methodology – The polar and the apolar extracts of two leaves and roots of each plant were analyzed by ¹H-NMR and GC-MS, respectively. Multivariate Data Analysis was used for a faster interpretation of data.

Results – The metabolic fingerprint of the Mediterranean plants, *Acanthus mollis*, *Dittrichia viscosa*, *Festuca drymeja*, *Fraxinus ornus*, *Fagus sylvatica*, *Hedera helix*, *Quercus ilex*, and *Typha latifolia*, showed a complex chemical composition, being specific for each species and plant tissue. Two alditols, mannitol and quercitol, were found in manna ash (*F. ornus*) and holm oak (*Q. ilex*) polar leaf extracts, respectively. The highest levels of aromatic compounds were found in *D. viscosa* and *T. latifolia*. Fatty acids were the predominant class of compounds in all apolar extracts under investigation. Triterpene were almost exclusively found in roots, except for holm oak, where they constitute 58% of total extract. Steroids were widespread in leaf extracts.

Conclusion – The major advantages of the proposed approach are versatility and rapidity, thus making it suitable for a fast comparison among species and plant tissue types.

50 **Introduction**

51 Higher plants produce a great diversity of chemical compounds, which have often a defensive role
52 and repellence against predators and microbial pathogens.¹⁻³ Moreover, environmental abiotic factors
53 like extreme temperature, light intensity and UV, shortage in water and minerals supply, as well
54 osmotic stress can cause the accumulation of specific metabolites in plants tissues.⁴

55 Exploitation of plant phytochemical diversity with antimicrobial activity in traditional
56 medicinal, as well for biological control in agriculture, has been the focus of an increasing number of
57 studies (Fabricant and Farnsworth, 2001, Koul and Dhaliwal, 2003).^{5,6} The study of medicinal plants
58 provides a scientific basis for the popular use against infectious diseases in the modern era.⁷ Although
59 thousands of plant species have been tested for antimicrobial properties, only a small fraction of the
60 estimated plant species has been investigated in their phytochemical composition,⁸ and, so the
61 majority of them has not been adequately evaluated. Moreover, the fractions of plant submitted to
62 biological or pharmacological screening are even smaller.^{9,10}

63 The phytochemical diversity of higher plants has been previously reviewed by examining
64 their involvement in constitutive¹¹ and inducible chemical defenses,¹² mechanisms of plant resistance
65 to biotic,¹³ as well to abiotic stresses,⁴ and fitness cost.¹⁴ The potential exploitation of such molecules
66 plant antimicrobial compounds has also been evaluated and thousands of diverse natural products,
67 involved in plant defense, have been identified including terpenoids, saponins, phenolics,
68 phenylpropanoids, pterocarpanes, stilbenes, alkaloids, glucosinolates, thiosulfinates and indoles.¹

69 Indeed, higher plants had enormous potential as sources for antimicrobial drugs with reference
70 to antibacterial and antifungal agents. However, the majority of previous studies have focused on a
71 single or few species, studying separately different plant parts (e.g. leaves, roots, flowers, seeds etc.)
72 However, no attempts targeted simultaneously several species and plant tissues to provide a
73 comprehensive description of plants metabolomics. Here, to overcome the limitations of previous
74 studies, we selected eight plant species from the Mediterranean biome (i.e. *Acanthus mollis* L.,
75 *Dittrichia viscosa* (L.) Greuter syn. *Inula viscosa* (L.) Aiton. *Fagus sylvatica* L., *Festuca drymeja*

76 Mert. et W. D. J. Koch, *Fraxinus ornus* L., *Hedera helix* L., *Quercus ilex* L. and *Typha latifolia* L.)
77 having different plant traits and being widely used as folk plants.¹⁵ We analyzed both leaf and root
78 organs using functionally complementary and powerful organic chemistry methods like Gas
79 Chromatography (GC-MS) and Nuclear Magnetic Resonance (NMR) Spectroscopy. Moreover, we
80 used a multivariate data analysis approach to manage the large amount of data produced by GC-MS
81 and NMR.

82 The main objectives of our study were:

- 83 (i) to describe the phytochemical composition and diversity of the eight plants;
84 (ii) to compare the leaf and root chemistry of selected Mediterranean plants.

85

86 **Experimental**

87

88 **Plant materials**

89 Eight Mediterranean plants were chosen for a metabolomic analysis through NMR and GC-MS
90 approaches. Leaves and roots of *A. mollis*, *D. viscosa*, *F. drymeja*, *F. ornus*, *F. sylvatica*, *H. helix*, *Q.*
91 *ilex*, and *T. latifolia* were collected in Cicerale (40°19' N, 15°07' E), an Italian municipality belonging
92 to the “Parco Nazionale del Cilento, Vallo di Diano e Alburni”, at an altitude of 250 m a.s.l.. The
93 study site has a typical Mediterranean climate with a mean annual temperature of 16.9 °C and 1,328
94 mm of annual rainfall well distributed in winter, spring and fall, but with a pronounced dry summer.

95

96 **Solvent and chemicals**

97 *n*-Hexane and methanol were obtained from Delchimica Scientific Laboratories (Naples, Italy).
98 Deuterium oxide (99,8 atom %D) was acquired from ARMAR Chemicals (Switzerland) and
99 chloroform-d (99,8 atom %D) and HCl- methanol solution 1.25 M were obtained from Sigma-Aldrich

100 (Steinheim, Germany). Dimethyl-4-silapentane sodium sulfonate (DSS) (Merck, Darmstadt,
101 Germany).

102 103 **Extraction procedure**

104 Leaves and roots of eight selected Mediterranean plants, collected in triplicates, were dried under
105 controlled temperature and powdered finely with a pestle and mortar. Four grams of each sample
106 were extracted with 20 mL of *n*-hexane under stirring for 1 h. Then, the apolar extract was filtered,
107 evaporated, and stored at 4 °C until analysis.

108 The plant material was further extracted with 20 mL MeOH/H₂O (6:4) solution under stirring
109 for 1 h, followed by centrifugation at 3000 rpm for five minutes at 25 °C. After separation, the polar
110 extract was collected, dried by a rotary evaporator, and stored in a refrigerator at 4 °C until analysis.
111 The following amounts of apolar extracts were obtained: *A. mollis* (33.7mg for leaves, 1.4 mg for
112 roots), *D. viscosa* (39.6 mg for leaves, 5.9 mg for roots), *F.drymeja* (9.4 mg for leaves, 1.8 mg for
113 roots), *F. ornus* (32.4 mg for leaves, 2.8 mg for roots), *F.sylvatica* (6.2 mg for leaves, 1.5 mg for
114 roots), *H. helix* (5.2 mg for leaves, 19.8 mg for roots), *Q. ilex* (6.7mg for leaves, 18.3 mg for roots),
115 *T. latyfolia* (5.4 mg for leaves, 1.7 mg for roots). The following amounts of polar extracts were
116 obtained: *A. mollis* (255.0 mg for leaves, 58.5 mg for roots), *D. viscosa* (468.0 mg for leaves, 285.7
117 mg for roots), *F.drymeja* (41.6 mg for leaves, 38.4 mg for roots), *F. ornus* (348.7mg for leaves, 95.4
118 mg for roots), *F.sylvatica* (20.1 mg for leaves, 60.0 mg for roots), *H. helix* (278.5 mg for leaves,
119 151.7 mg for roots), *Q. ilex* (165.4 mg for leaves, 340.0 mg for roots), *T. latyfolia* (150.0 mg for
120 leaves, 54.5 mg for roots).

121 All samples were analyzed in triplicate to ensure their reproducibility. Apolar extracts was
122 analyzed by GC-MS, while polar extracts was analyzed by ¹H-NMR.

123 124 **NMR Experiment**

125 An aliquot (10 mg) of each dried polar sample was solubilized in 600 µl of deuterium oxide (99.9%
126 D₂O) and transferred into a 5 mm NMR tube. Dimethyl-4-silapentane sodium sulfonate (DSS)

(Merck, Darmstadt, Germany), added at a concentration of 0.2 mg/mL, was used as an internal standard. The NMR spectra were recorded at 298 K on a Varian Unity Inova spectrometer operating at 500 MHz. The ¹H-NMR experiments were performed with 128 transients and 16K complex data point. The recycle time was set to 5 s, and a 45° pulse angle was used. Chemical shifts were referred to DSS signal (δ 0.00 ppm). All spectra were processed using iNMR program (www.inmr.net), phased and baseline corrected. In total, 48 spectra (16 plant population × 3 replicates) were acquired. Quantification was performed by signal integration relative to the internal standard, DSS. The region of the solvent peaks was excluded from the analysis. Spectral peak assignments of the detected compounds were obtained based on pure standards purchased by Sigma-Aldrich, and on combined comparison with data reported in the literature and in Human Metabolome Database (HMDB).^{16,17} All spectra were manually phased, and baseline corrected.

Gas chromatography–mass spectrometry

Apolar extracts were derivatized as methyl esters before analysis by GC-MS. For this purpose, an aliquot of each apolar extract (0.5 mg) was transferred into a vial and dissolved in 1 ml of a solution of MeOH: HCl 1.25 N. The vials were vortexed and left at 50 °C overnight, then they were neutralized with NaOH 1N, dried under nitrogen, solubilized in n-hexane and analyzed by GC-MS. Chromatographic conditions are those described by de Falco *et al.*^{18,19} One µl of derivatized samples were injected in a pulsed splitless mode into an Agilent-7820A GC system with 5977E MSD operating in EI mode at 70 eV. The system was equipped with a 30 m × 0.25 mm id fused-silica capillary column with 0.25 µm HP-5MS stationary phase (Agilent Technologies, UK). The injection temperature was set at 270 °C. Helium was used as carrier gas at a constant flow rate of 1 ml/min. Separation of the nonpolar extract was achieved using a temperature program of 80 °C for 1 min, then ramped at 10 °C/min to 320 °C and held for 1 min. Both chromatograms and mass spectra were evaluated using the MassHunter Qualitative Analysis B.07.00 (Agilent Technologies, CA, USA). Mass spectra of all detected compounds were compared with standard compounds and with spectra

153 in National Institute of Standard and Technologies library NIST MS search 2.2. Data was processed
154 with the AMDIS (Agilent Technologies, CA, USA) software to deconvolute co-eluting peaks.
155 Compounds were selected from the NIST list on the basis of the probability of identification (> 90%)
156 and by comparison with literature data. The relative amounts of separated metabolites were calculated
157 from Total Ion Chromatography (TIC) by the computerized integrator.

158

159 **Multivariate data analysis**

160 Resulting dataset from ¹H-NMR and GC-MS analysis was examined through Multivariate approach
161 in order to obtain information of its underlying structure and the effect of multiple variable on the
162 chemical differentiations between the plant species and plant tissues object of the study. Previous to
163 apply multivariate approach each dataset was normalized. Particularly for NMR analysis, data were
164 normalized to total area to minimize small differences and subsequently mean-centered. In detail, for
165 GC-MS analytical method data list of compounds was aligned according by time of retention time
166 between samples. Aligned data was utilized to build a double entry data matrix, with plant species
167 and portion as cases in column and identified GC-MS observed metabolite as variable in row. When
168 metabolite was not detected for a plant species or portion, 0 value was assigned.

169 For ¹H-NMR, the description of statistical analyses refers to range scaled data, in order to
170 preserve experimental biological information. Total dataset was plotted according to PCA, in order
171 to explain main chemical species producing differentiations among plant and root extracts. Given the
172 high number of resonance regions and the unbalanced presence of chemical classes that are
173 constitutively more produced with respect to other, we perform three additional PCA on different
174 resonance regions. Resonance regions were clustered according to common chemical classes as
175 described following: i) Aromatic/ phenolic compounds regions (from δ 10.5 to 5.5); ii) Carbohydrates
176 regions (from δ 5.5 to 3.0) iii) Aliphatic regions (from δ 3.0 to 0.5).

177 Data ordination and normalization was performed by means of Excel software, while
178 Multivariate analysis and plotting was performed in Statistica 10 software (StatSoft, Inc., Tulsa, OK).

179

180 **Results and discussion**

181

182 **Metabolite profiling of polar extracts**

183 An integrate spectroscopic approach combined with multivariate data analysis was applied on eight
184 Mediterranean plants. The metabolic profile of leaves and roots was obtained to comprehensively
185 evaluate the metabolome of each species and how its chemical composition was distributed in two
186 compartments of each plant species. On the basis of our previous experience,^{16, 17} the polar extracts
187 were analyzed by NMR analysis, while the apolar extracts were investigated through GC-MS,
188 because of the strong overlapping of the methylene signals in the ¹H-NMR spectra.

189 Each polar extract showed a very intricate profile, with free aliphatic and aromatic amino acids,
190 carbohydrates, organic acids and aromatic compounds; the qualitative and quantitative metabolite
191 profile was peculiar of each analyzed species (Figure 1).

192 For more convenient data interpretation, the ¹H-NMR spectra were divided in three regions:
193 the aliphatic region between 0.5-3.10 ppm, the sugar region between 3.10-5.50 ppm and the aromatic
194 region ranging from 5.50 to 8.5 ppm (Figure 2). The aliphatic region contained signal related to amino
195 acids and organic acids. Diagnostic methyl doublets typical of isoleucine (Ile) and valine (Val),
196 resonated at 0.91 ppm and 1.01 ppm, respectively, and the methyl triplet of leucine (Leu) at 0.95 ppm
197 allowed their qualitative and a quantitative assignment. Moreover, doublets at δ 1.46 (J 7.0 Hz) and
198 δ 1.32 were associated to alanine (Ala) and threonine (Thr), respectively (Table S1). The typical
199 region of methylene groups closes to a carbonyl group in ¹H-NMR spectra showed a triplet at 2.98
200 ppm attributed to the γ -methylene protons of γ -amino butyric acid (GABA), as well as two double
201 doublets at 2.84 and 2.94 ppm, corresponding to the diastereotopic hydrogens of asparagine (Asn). A
202 mention is due to proline (Pro), whose recognized has been obtained by three multiplets at δ 1.99,
203 2.06 and 2.34, and to glutamic acid (Glu) with the typical multiplet signals at δ 2.05, 2.10 and 2.36.
204 Pro and Glu were not always present in the studied species, but when they occurred in the plant were

205 present at reasonable amounts, although it was not easy to distinguish between them (Table S1). All
206 monosaccharides and alditols were quantified by integrating the signals indicated in Table S1. Finally,
207 the aromatic region was selected from 5.51 to 8.50, excluding three multiplet signals at 7.32, 7.36
208 and 7.40 ppm corresponding to phenylalanine (Phe), and two doublets at 6.80 and 7.12 ppm,
209 corresponding to tyrosine (Tyr). Some aromatic signals were determined, as chlorogenic acid (CA)
210 (Table S1).

211 The results showed that all analyzed samples have carbohydrates as major metabolites. In
212 detail, the analysis of the leaves indicated *F. ornus* and in *Q. ilex* to contain a rather high content of
213 monosaccharides, due to the presence of additional alditols. In the ¹H-NMR spectra of *F. ornus*
214 leaves, the signals of mannitol were easily recognized by the presence of two coupled double doublets
215 at δ 3.66 and 3.85, a double triplet at δ 3.75 and a doublet at δ 3.79. Mannitol is the major component
216 of manna, which is produced from *Fraxinus* sp. especially under stress conditions.²⁰ In our study
217 mannitol alone represented 45.1% in weight of the total metabolome of *F. ornus*. The holm oak (*Q.*
218 *ilex*) contained two metabolites deriving from the shikimic acid pathway,²¹ quercitol and quinic acid
219 (QA), whose signals resonated mostly in the sugar region. This is probably the reason for the high
220 sugar content found for this species. It has been reported that QA and quercitol are the most abundant
221 metabolites in *Q. ilex* and in other species of *Quercus*;^{22,23} their production is a reaction to biotic and
222 osmotic stress.^{24,25} Quantitative determination of QA and quercitol was not easy due to their nearness
223 in the ¹H-NMR spectra; to avoid any kind of overlapping, we choose to integrate the signal at δ 1.81
224 for quercitol and the signal at δ 1.87 for QA. In this way, we were able to quantitate quercitol and QA
225 which represented 18.9% and 13.9% of the all holm oak leaves polar extract, respectively. On the
226 contrary, *D. viscosa* had the lowest amount of carbohydrates (10.4%) (Figure 3, Table S1 and S2).

227 The organic acid total content was almost the same in all analyzed leaves (~10%) with the
228 exception of *T. latifolia*, *H. helix*, and *Q. ilex*. The former contained the lowest content of organic
229 acid (2.6%) among the analyzed leaves. The latter present the highest content, reaching respectively
230 23.2% and 21.0%, due to the presence of QA, absent in the other analyzed plants (Figure 3, Table

231 S2). Moreover, the leaves of *A. mollis* showed a high amount of betaine, recognizable from a singlet
232 at 3.25 ppm (Table S1).

233 The leaves of *D. viscosa*, *F. sylvatica*, and *F. drymeja* showed a high percentage of amino
234 acids, which was partially due to the high values of glutamic acid (31.9%, 24.8% and 20.9%,
235 respectively) (Figure 3). Aromatic compounds were particularly abundant in *D. viscosa* and *T.*
236 *latifolia* leaves, followed by *F. sylvatica*, *F. ornus* and *A. mollis*. The lowest content of aromatic
237 compounds was found in *Q. ilex* leaves (Figure 3).

238 Concerning root tissue, the carbohydrate content of the analyzed species was generally around
239 50% of the comprehensive metabolite content of polar extract, with some notable exceptions. *A.*
240 *mollis* extract had 80.6% of sugar content, due to the presence of several sugar residues, from which
241 raffinose (Raff) was predominant with 36.5% of total extract. *F. ornus* contained 62.7% of sugars,
242 the most abundant being sucrose (Sucr). Moreover, particularly low is the content of amino acids and
243 organic acids in these species. On the contrary, *D. viscosa* and *T. latifolia* had the lowest percentage
244 of carbohydrate content and the highest percentage of aromatic compounds (Figure 3).

245

246 **Metabolite profiling of apolar extracts**

247 NMR spectra of the apolar extracts of each plant showed a chemical composition almost exclusively
248 amenable to fatty acids. The overlapping of some signals, such as the methylene groups close to a
249 carbonyl group, did not allow to distinguish and quantify all fatty acids of apolar extracts. For this
250 reason, we used gas chromatography coupled to mass spectrometry (GC-MS) with increasing
251 separation capability, allowing to determine the quali-quantitative profile of the studied
252 Mediterranean plants. In this way, it was possible to identify single fatty acids on the basis of their
253 molecular weight.

254 The GC-MS data of the species under investigation allowed to characterize 60 metabolites,
255 belonging to several classes of organic compounds. The count of metabolites extracted was higher in
256 the root extracts, reaching the maximum value in *F. ornus*, which also contained the highest

257 variability of organic compounds (Figure 4). Fatty acids were ubiquitous metabolites in all analyzed
258 samples at variable percentage, ranging from 24.5% in *D. viscosa* to 100% in *T. latifolia*. Moreover,
259 the fatty acid profile with relative amount depend on the species analyzed with palmitic and oleic
260 acids always present (Table S3). In particular, palmitic acid showed an elevate concentration range
261 from a maximum value of 52.7% in *A. mollis* roots to a minimum value of 1.0% in *T. latifolia* leaves.
262 Similarly, cerotic acid (C26:0) has been determined in most analyzed plants at very different
263 concentration (44.6% in *A. mollis* leaves, 30.1% in *T. latifolia* leaves and 1.9% in roots) while in
264 others (e.g. *H. helix*) it was absent. Unsaturated fatty acids were determined in the roots of *F. ornus*
265 and *H. helix* and identified as C16:1, C18:2, C18:1 and C20:1, the last found only in manna ash (Table
266 S3).

267 The apolar extracts of leaves contained sesquiterpenes and alkanes in addition to fatty acids.
268 The detected heptacosane, nonacosane and heneicosane alkanes were determined in *F. ornus* and *F.*
269 *sylvatica*, representing 48.9% and 43.9% of total extract, and in *D. viscosa*, where they were less
270 abundant (Table S3). Sesquiterpenes were almost exclusively found in *D. viscosa*, where they
271 represented 59.4% of the total extract and in the roots of *A. mollis* and *F. ornus* roots, present at
272 smaller amounts (Table S3). Triterpenes were widespread in all studied roots extracts, although holm
273 oak leaves contained a high level of these organic compounds (57.9% of total extract). Triterpenes
274 in roots were present at variable concentrations and showed the highest content in *A. mollis*, and *F.*
275 *drymeja*, *D. viscosa*, *Q. ilex*, and *T. latifolia*, the latter species containing 38.3% and 28.4% of lupan-
276 3-one and friedelan 3-one, respectively (Table S3).

277 Finally, steroids were found in *A. mollis* (16.7%), *F. drymeja* (41.2%) and *H. helix* (27.3%)
278 leaves, and *F. sylvatica* (14.1%) and *Q. ilex* roots (41.0%), where the most representative steroid was
279 3b,5a,6b-Cholestanetriol (Table S3).

280

281 **Multivariate Data Analysis**

282 Coupled with metabolic profiling, the multivariate approach ordinated plant species according to their
283 respective metabolic characteristics. Through PCA we obtained a general view of the underling
284 structure of the data. The principal components were displayed as a set of scores (PC), which
285 highlights clustering or outliers, and a set of loadings (p), which emphasizes the influence of input
286 variables on PC. The multivariate methodology was applied both for polar and apolar extracts and,
287 particularly for data originated from NMR, was tested for the ordination of plant according to
288 chemical characteristics present in three different regions putatively assigned to aromatic/ phenolic
289 compounds (from δ 10.5 to 5.5), carbohydrates regions (from δ 5.5 to 3.0), and aliphatic compounds
290 (from δ 3.0 to 0.5).

291 In PCA performed for the totality of the regions from ^1H -NMR spectra, the first 2 components
292 explains the 71.4% of the variance among the samples (PC1 65.6 and PC2 5.8%). Results are showed
293 in figure 4A and 4B for loadings and score plots, respectively. In a general view, we observed a
294 marked ordination of loadings values according to the respective plant species. Inversely, metabolic
295 profile of the plants does not discriminate among plant portion from which metabolite was extracted.
296 The general variation among plant species was triggered by carbohydrates, while aliphatic and
297 aromatic/phenolics regions has a decreased discriminant power. This is likely explained by the normal
298 attitude to accumulate carbohydrates as nutrient source from photosynthetic pathways. ²⁶Given this,
299 carbohydrates mediate the unidirectional disposition of the samples in its correspondent area showing
300 a generalized positive association of all the samples with carbohydrate signals. However, peculiar
301 number of specific metabolites generate distinctive disposition among plant species. For instance, *D.*
302 *viscosa* leaves extracts is characterized by the presence of betaine, that also appears to be responsible
303 of the separation of *A. mollis* leaves extracts from the other plant extracts. So far, mannitol signals
304 are majorly associated to the well-known manna producer species *F. ornus* and the quercitol to the
305 oak *Q. ilex*. In both the cases, the metabolite disposition appears to be few distinctive by the point of
306 view of the plant organs in which the metabolite was extracted. More generally, *H. helix*, *F. ornus*,
307 *F. drymeja*, *F. sylvatica*, *D. viscosa* for roots and *F. ornus*, *F. drymeja*, *F. sylvatica* for leaves

308 associate with the aforementioned mannitol and glutamic acid, sucrose, fructose, shikimic acid and
309 raffinose.

310 To avoid the hiding action of carbohydrates on the other spectral regions we analyzed these
311 in separate way with the same multivariate approach. Figure 6A and 6C showed the PCA ordination
312 of different extracts according to their chemical composition. The PCA explained totally the 62.0%
313 of the variance in the sample (PC1 51.7% and PC2 10.3%). In this case, we observed a marked
314 differentiation of *A. mollis* and *H. helix* root and leaf extracts with respect to other species mainly
315 operated by the higher content of fumaric acid. Intermediate position is instead acquired by *F.*
316 *drymeja* leaves extract associated by higher content of tyrosine. The other species aggregates in same
317 directional ordination that is given by the similarity of the spectral regions between 6.96 and 6.83
318 ppm. For those regions, we unassigned the signals given the high level of uncertainty. In addition,
319 residuals signals from carbohydrates and amino acidic compounds interfere in the interpretation of
320 the spectra limiting our multivariate approach. For the PCA of carbohydrate region (Figure 6C and
321 6D), is observed a specular disposition of the data to those of the comprehensive PCA (Figure 5A
322 and 5C). Indeed, the PCA of carbohydrate region explain 72.6% of the variance with respect to the
323 71.4% explained by the PCA of the overall dataset. For the PCA of the aliphatic region, lower level
324 of explained variance was reported (Figure 6E and 6F). Nonetheless, *A. mollis* and *Q. ilex* leaves and
325 *D. viscosa* roots differentiate for quinic acid and proline contents rather than other species that
326 differentiate for the contents of threonine and signals of rhamnose, acetic acid/GABA and residual
327 signals from polar portion of fatty acids.

328 PCA from polar extract showed lower level of explained variance (Total of PC 46.6%, Figure
329 5C and 5D). However, we observed a clearer segregation of root extract with respect leaves extracts,
330 that make only exception for *F. ornus* and *Q. ilex* leaves. Interestingly, *Q. ilex* root extracts dispose
331 with leaves extract of other species in complete opposite way respect the general behaviors in the
332 samples. So far, root chemical differentiation from apolar extracts is given by the presence of palmitic
333 acid, 2,6,6,9,2',6',6',9'-Octamethyl-[8,8']bi[tricyclo[5.4.0.0(2,9)]undecyl], linoleic acid and behenic

334 acids. Oppositely, leaves are higher in contents of cerotic acid, montanic acid, eptacosane, lignoceric
335 acid, cholestane and arachidic acid.

336

337 **Conclusions**

338 The large number of biological questions on plant metabolism requires that answers should be sought
339 using the most versatile techniques available. Our goal was the development of a method capable to
340 detect most classes of organic compounds.²⁷ Unfortunately, the huge variety of chemical compounds
341 found in plants did not allow the use of a single technique to this aim.

342 The approach we developed to evaluate the chemical composition of eight Mediterranean plant
343 species was based on spectroscopic techniques and multivariate data analysis. It was appropriate for
344 a fast and comprehensive analysis of primary and secondary metabolites, allowing the determination
345 of a metabolomic fingerprint of each species and the evaluation of the different distribution of the
346 metabolites in two parts (leaves and roots) of the plant.

347 *Q. ilex* and *F. ornus* contained large amounts of specific metabolites, quinic acid, quercitol and
348 mannitol, usually produced from plants during stress conditions. Besides being involved in osmotic
349 stress, quercitol has been recently used as a building block in the synthetic strategy for antidiabetic
350 compounds.²⁸ *D. viscosa* was characterized by a high content of aromatic compounds at the expense
351 of carbohydrate production in the polar fraction, and a considerable content of sesquiterpenes in the
352 apolar fraction. The separation of *A. mollis* from the other species was due to the presence of betaine
353 and sucrose in leaves and raffinose in roots.

354 Moreover, fatty acids were present in all the analyzed species, especially in *T. latifolia* leaves and in
355 *A. mollis* and *F. drymeja* roots. Fatty acids were almost exclusively present as saturated fatty acids.

356 In conclusion, the approach we developed proved to be suitable for a rapid investigation of different
357 plant species containing a wide range of organic molecules.

358

359

360 **Acknowledgements**

361 NMR spectra and GC-MS analyses were recorded at the “Laboratorio di Analisi strumentale”,
362 Dipartimento di Farmacia, Università di Napoli Federico II; the assistance of Paolo Luciano is
363 gratefully acknowledged.

364 **Notes**

365 The authors declare no competing financial interest.

366

References

1. Dixon RA. Natural products and plant disease resistance. *Nature* 2001; 411: 843–847.
2. Hartmann T .The lost of origin of chemical ecology in the late 19th century. *Proc. Natl. Acad. Sci. USA* 2008; 105: 4541–4546.
3. Springob K, Kutchan TM. Introduction to the different classes of natural products. In: *Plant-derived Natural Products*. Springer Science; 2009. 3 p.
4. Akula R, Ravishankar, GA. Influence of abiotic stress signals on secondary metabolites in plants. *Plant Signal Behav.* 2011; 6(11): 1720-1731.
5. Fabricant DS, Farnsworth NR. The value of plants used in traditional medicine for drug discovery. *Environ Health Perspect* 2001; 109 (suppl 1): 69-75.
6. Koul O, Dhaliwal GS. *Phytochemical biopesticides*. CRC Press 2003 (Vol. 1).
7. Gibbons S. Plants as a source of bacterial resistance modulators and anti-infective agents *Phytochem Rev* 2005; 4(1): 63-78.
8. Rates SM. Plants as source of drugs. *Toxicon* 2001; 39: 603-613.
9. Hamburger M, Hostettmann K. Bioactivity in plants: the link between phytochemistry and medicine. *Phytochem* 1991; 30 (12): 3864-3874.
10. Hostettmann K, Potterat O, Wolfender J L. The potential of higher plants as a source of new drugs. *Chimia* 1998; 52(1-2): 10-17.

-
11. Wittstock U, Gershenzon J Constitutive plant toxins and their role in defense against herbivores and pathogens. *Curr Opin Plant Biol* 2002; 5:1–8
 12. Hammerschmidt R. Phytoalexins: what have we learned after 60 years? *Annu Rev Phytopathol* 1999; 37: 285–306
 13. Morrissey JP, Osbourn AE. Fungal resistance to plant antibiotics as a mechanism of pathogenesis. *Microbiol Mol Biol R* 1999; 63: 708–724.
 14. Heil M. Ecological costs of induced resistance. *Curr Opin Plant Biol* 2002; 5: 1–6
 15. Motti R, Ippolito F, Bonanomi G. Folk phytotherapy in paediatric health care in Central and Southern Italy: a review. *Human Ecology* 2018; 46(4): 573-585.
 16. de Falco B, Incerti G, Pepe R, Amato M, Lanzotti V. Metabolomic fingerprinting of romaneschi globe artichokes by nmr spectroscopy and multivariate data analysis: metabolomics of artichokes by nmr and chemometrics. *Phytochem Anal* 2016; 27: 304–314.
 17. de Falco B, Incerti G, Bochicchio R, Phillips, TD, Amato M, Lanzotti, V. Metabolomic analysis of *Salvia hispanica* seeds using NMR spectroscopy and multivariate data analysis. *Ind Crops Prod* 2017; 99: 86-96.
 18. de Falco B, Fiore A, Rossi R, Amato M, Lanzotti V. Metabolomics driven analysis by UAEGC-MS and antioxidant activity of chia (*Salvia hispanica* L.) commercial and mutant seeds. *Food Chem* 2018; 254: 137–143.
 19. de Falco B, Fiore A, Bochicchio R, Amato M, Lanzotti V. Metabolomic analysis by UAE-GC MS and antioxidant activity of *Salvia hispanica* (L.) seeds grown under different irrigation regimes. *Ind Crops Prod* 2018; 112: 584–592.
 20. Stoop JMH, Williamson JD, Mason P. Mannitol metabolism in plants: a method for coping with stress. *Trends Plant Sci* 1996; 1(5): 139-144.
 21. Wilson DJ, Patton S, Florova G, Hale V, Reynolds KA (1998) The shikimic acid pathway and polyketide biosynthesis. *J Ind Microbiol Biotechnol* 1998; 20: 299–303.

-
22. Sardans J, Gargallo-Garriga A, Pérez-Trujillo M, Parella TJ, Seco R, Filella I, Penuelas J.). Metabolic responses of *Quercus ilex* seedlings to wounding analysed with nuclear magnetic resonance profiling. *Plant Biol.* 2014; 16(2): 395-403.
23. Passarinho JAP, Lamosa P, Baeta JP, Santos H, Ricardo CPP. Annual changes in the concentration of minerals and organic compounds of *Quercus suber* leaves. *Physiol Plant.* 2006; 127: 100–110.
24. Koskimäki JJ, Hokkanen J, Jaakola L, Suorsa M, Tolonen A, Mattila S, Pirttilä AM, Hohtola A. Flavonoid biosynthesis and degradation play a role in early defence responses of bilberry (*Vaccinium myrtillus*) against biotic stress. *Eur J Plant Pathol* 2009; 125: 629–640.
25. Adams MA, Richter A, Hill AK, Colmer TD. Salt tolerance in *Eucalyptus* spp.: identity and response of putative osmolytes. *Plant Cell Environ* 2005; 28: 772–787.
26. Heldt HW, Piechulla B, Heldt F. *Plant Biochemistry*. Translation of the 4th German edition. 2011
27. de Falco B, Lanzotti V. NMR spectroscopy and mass spectrometry in metabolomics analysis of *Salvia*. *Phytochem Rev* 2018; 17(5): 951-972.
28. Worawalai W, Sompornpisut P, Wacharasindhu S, Phuwapraisirisan P. Quercitol: from a taxonomic marker of the genus *Quercus* to a versatile chiral building block of antidiabetic agents. *J Agric Food Chem.* 2018; 66(23): 5741-5745.

367 **FIGURE CAPTIONS**
368

369 **Figure 1.** ^1H -NMR at 500 MHz in D_2O of leaves (L) and roots (R) of Mediterranean species.

370

371 **Figure 2.** ^1H -NMR spectra at 500 MHz in D_2O of *A. mollis* and *Q. ilex* leaves (L) and roots (R).

372

373 **Figure 3.** Heat-map of the relative concentration (%) of metabolites in leave and root polar extract of
374 each species.

375

376 **Figure 4.** a) Total count of organic compounds in apolar leaf (left) and root (right) extracts, and; b)
377 relative concentration (%) for class of organic compounds in apolar leaf (left) and root (right) extracts.

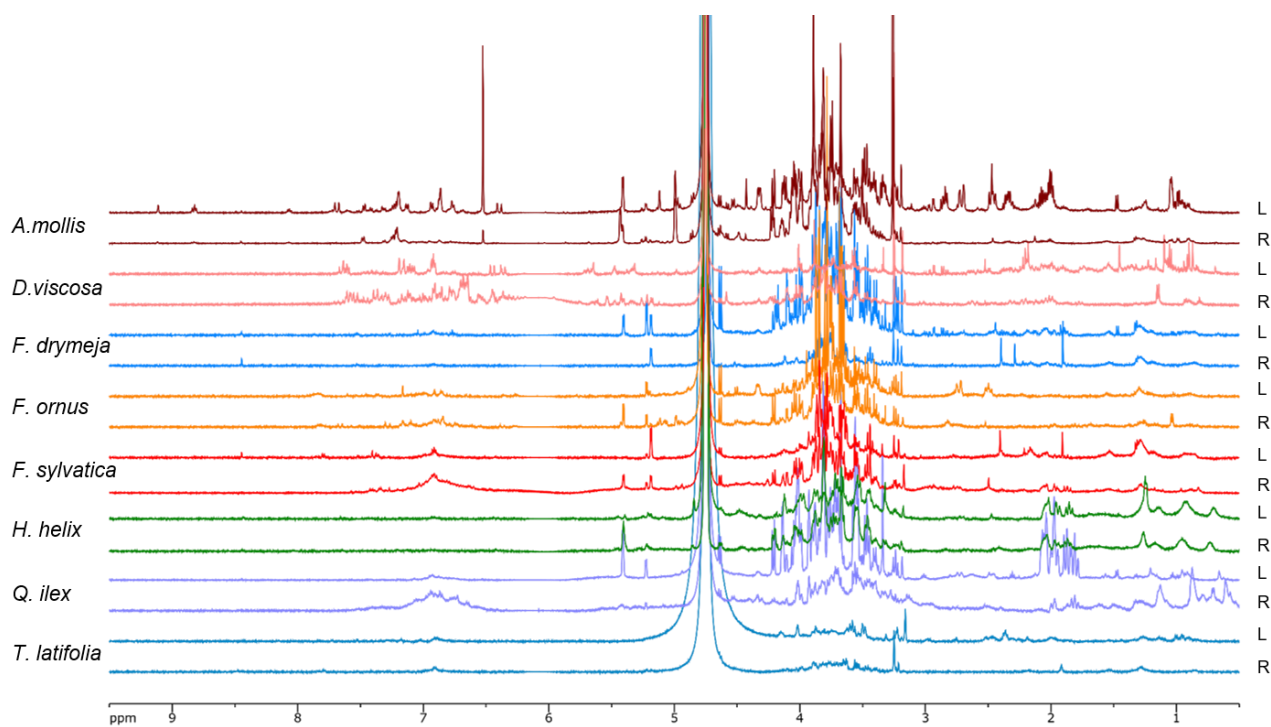
378

379 **Figure 5.** Principal component analysis (PCA) ordination of eight Mediterranean plant leaves and
380 roots based on ^1H -NMR resonance spectra from polar (A and B) and apolar (B and C) fractions. A
381 and C: variable loadings; B and D: factorial scores of resonance intervals of 0.01 ppm and retention
382 time value. Explained variance of principal components is reported on the axis labels. Plants in
383 loading plots are numbered as: 1. *A. mollis*, 2. *D. viscosa*, 3. *F. drimejia*, 4. *F. ornus*, 5. *F. sylvatica*,
384 6. *H. helix*, 7. *Q. ilex*, 8. *T. latifolia*.

385

386 **Figure 6.** Principal component analysis (PCA) ordination of ^1H -NMR resonance intervals: (A and B)
387 from δ 10.5 to 5.5; (C and D) from δ 5.5 to 3.0; (E and F) from δ 3.0 to 0.5. Left: variable loadings;
388 right: factorial scores of resonance intervals of 0.01 ppm. Explained variance of principal components
389 is reported on the axis labels. Plants in loading plots are numbered as: 1. *A. mollis*, 2. *D. viscosa*, 3.
390 *F. drimejia*, 4. *F. ornus*, 5. *F. sylvatica*, 6. *H. helix*, 7. *Q. ilex*, 8. *T. latifolia*.

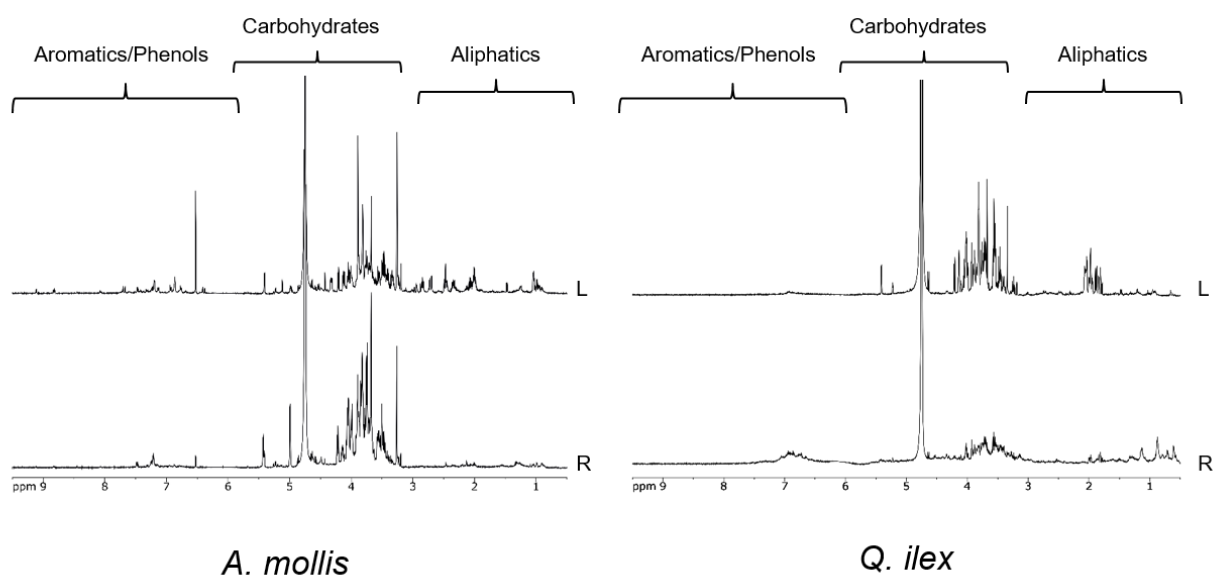
391



392

393 Figure 1. ^1H -NMR at 500 MHz in D_2O of leaves (L) and roots (R) of Mediterranean species.

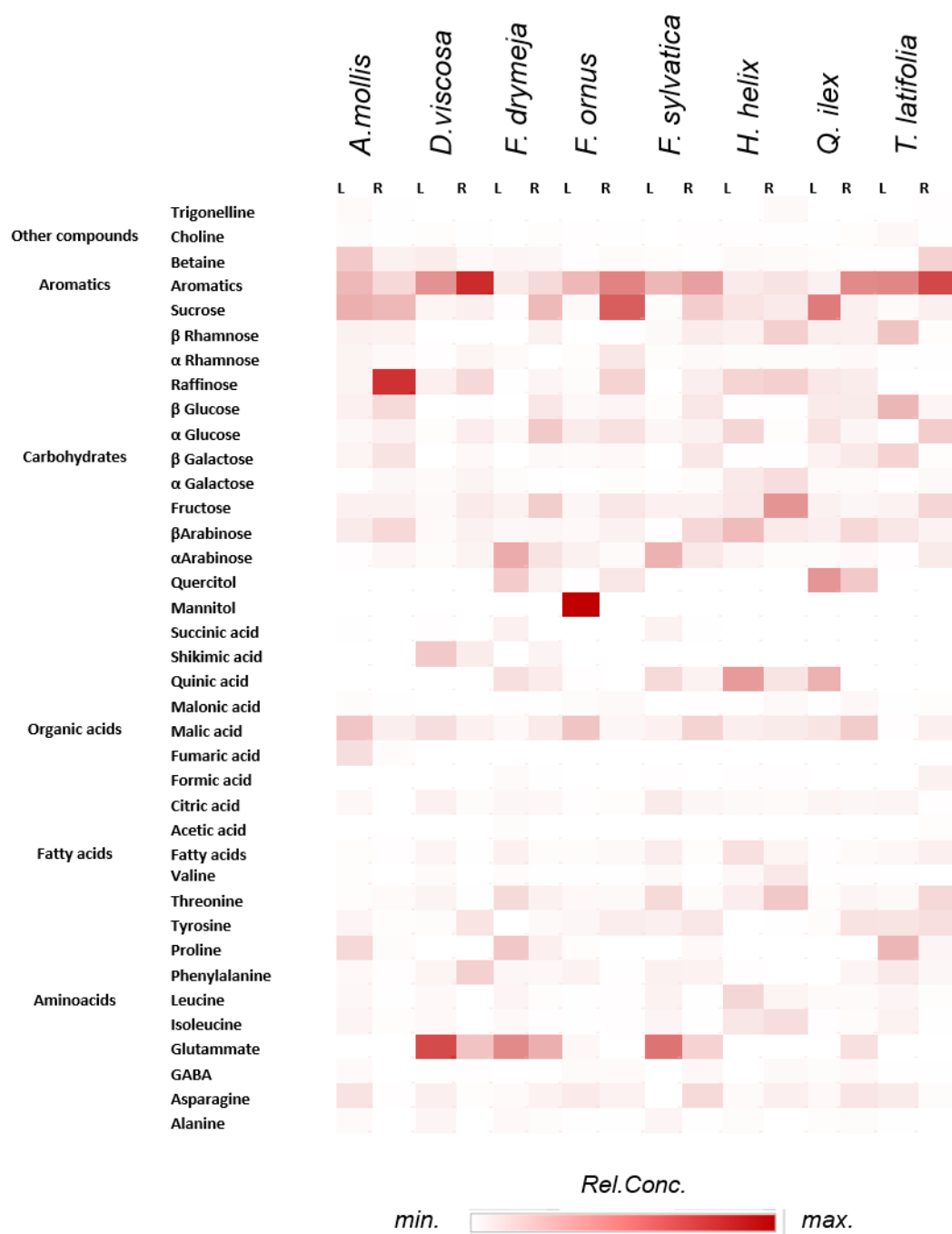
394



395

396 Figure 2. ^1H -NMR spectra at 500 MHz in D_2O of *A. mollis* and *Q. ilex* leaves (L) and roots (R).

397



398

399 Figure 3. Heat-map of the relative concentration (%) of metabolites in leave and root polar extract of
 400 each species.

401

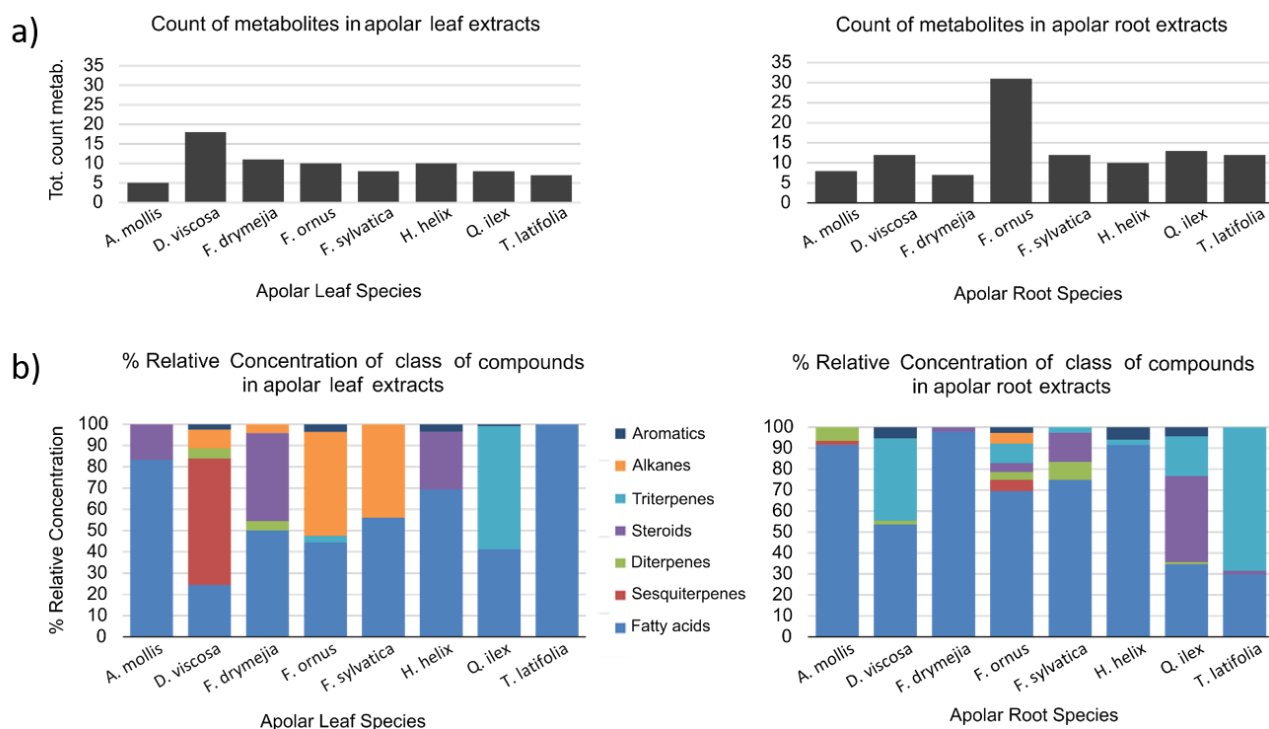


Figure 4. a) Total count of organic compounds in apolar leaf (left) and root (right) extracts; b) Relative concentration (%) for class of organic compounds in apolar leaf (left) and root (right) extracts

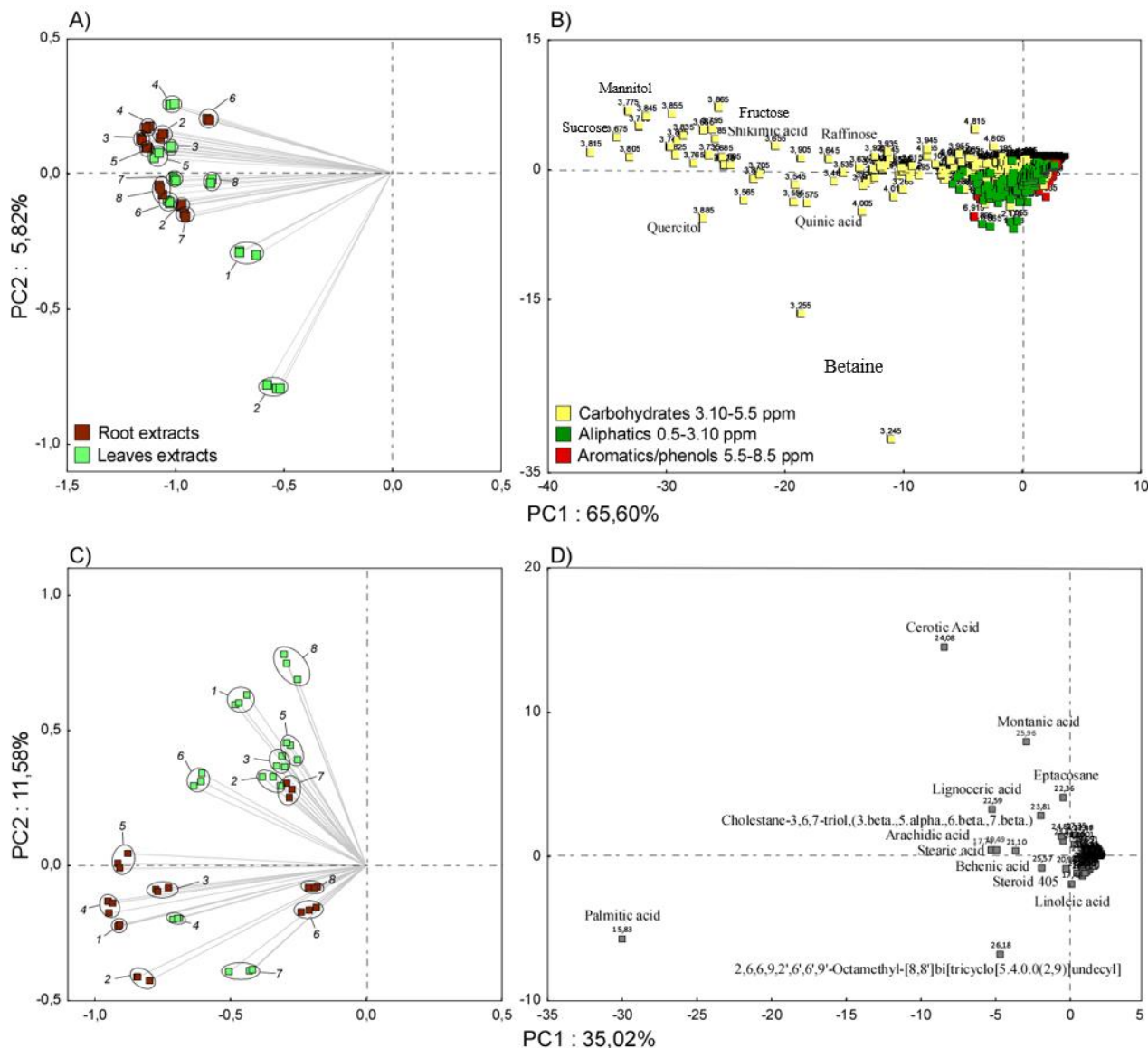
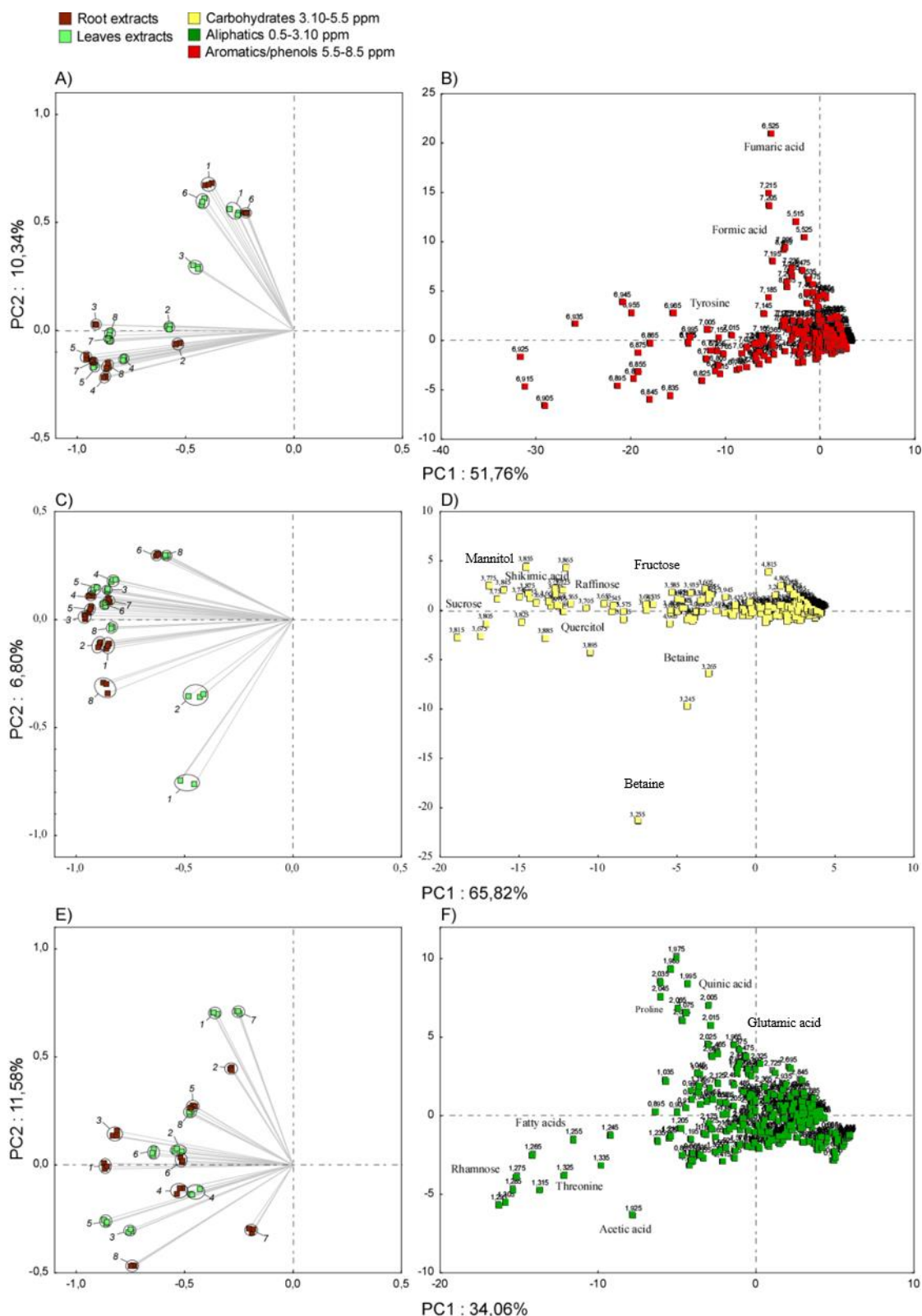


Figure 5. Principal component analysis (PCA) ordination of eight Mediterranean plant leaves and roots based on ^1H -NMR resonance spectra from polar (A and B) and apolar (B and C) fractions. A and C: variable loadings; B and D: factorial scores of resonance intervals of 0.01 ppm and retention time value. Explained variance of principal components is reported on the axis labels. Plants in loading plots are numbered as: 1. *A. mollis*, 2. *D. viscosa*, 3. *F. drimeja*, 4. *F. ornus*, 5. *F. sylvatica*, 6. *H. helix*, 7. *Q. ilex*, 8. *T. latifolia*.



415

416 Figure 6. Principal component analysis (PCA) ordination of ^1H -NMR resonance intervals: (A and B)
 417 from δ 10.5 to 5.5; (C and D) from δ 5.5 to 3.0; (E and F) from δ 3.0 to 0.5. Left: variable loadings;
 418 right: factorial scores of resonance intervals of 0.01 ppm. Explained variance of principal components
 419 is reported on the axis labels. Plants in loading plots are numbered as: 1. *A. mollis*, 2. *D. viscosa*, 3.
 420 *F. drimeja*, 4. *F. ornus*, 5. *F. sylvatica*, 6. *H. helix*, 7. *Q. ilex*, 8. *T. latifolia*.